Oxford elbow scores in an asymptomatic population

PM Guyver, AE Cattell, MJ Hall, MD Brinsden

Plymouth Hospitals NHS Trust, UK

ABSTRACT

INTRODUCTION Patient reported outcome measures (PROMs) are being used increasingly to assess the quality of healthcare delivery in the UK. It is important when using PROMs to know the score of the background population against which any clinical intervention may be benchmarked. The purpose of this study was to measure an elbow specific PROM for the population of the South West Peninsula.

METHODS A cross-sectional survey was undertaken of patients and healthcare professionals from the South West Peninsula population. Participants were asked to complete a simple demographic questionnaire and an Oxford elbow score for each elbow. Respondents with a history of elbow surgery, elbow injury, chronic elbow problems or an incomplete dataset were excluded from the study.

RESULTS A total of 1,765 respondents (3,530 elbows) completed the survey but 567 questionnaires were excluded due to incomplete datasets, pre-existing elbow pathology or age criteria, leaving 1,198 individuals (2,396 elbows) for analysis. The median scores for each decade group were 48, with mean scores ranging from 46.74 to 47.94. There was no significant clinical difference in the scores for age, sex or hand dominance.

CONCLUSIONS When using the Oxford elbow score to assess outcomes after surgery, a normal score should be used as the benchmark. This benchmark is independent of age, sex and hand dominance.

KEYWORDS

Oxford elbow score – Patient reported outcome measures – Background population

Accepted 19 April 2013

CORRESPONDENCE TO

Andrew Cattell, Orthopaedic Registrar, Royal Cornwall Hospital, Truro, Cornwall TR1 3LJ, UK
E: cattellandrew@yahoo.co.uk

The Oxford elbow score (OES) was developed in response to reports of a lack of an appropriate standardised method of outcome assessment in elbow surgery, leading to inadequate evaluation of many elbow operations.¹ The OES was designed specifically to be a patient focused outcome measure so as to be independent from the operating surgeon and therefore minimise the risk of bias.² Research showing that assessments of health status made by healthcare professionals can differ from those made by the patient³ reinforced the argument that well developed patient reported outcome measures (PROMs) should be sought to provide an objective measure of a subjective construct.⁴ The results of the OES development study demonstrated that it was reliable, valid and acceptable to patients, with a high rate of completion. Moreover, it was a useful method of assessing the outcome of elbow surgery from the patient's perspective.⁵

The OES has 12 items (questions) with 5 response options each. Each item response is scored as 0 to 4, with 0 representing greater severity. Underlying the 12 items are 3 domains (subscales): elbow pain, elbow function and social-psychological effects. Scores for each domain are

calculated as the sum of each individual item score within that domain. This gives a score range of 0–16 for each domain and 0–48 overall, with 0 indicating the worst elbow score and 48 as a 'normal' elbow score. Individual domain scores can be further analysed by being converted to a metric score of 0–100 (lower scores representing greater severity) using the equation:

$$\frac{100}{\text{Maximum possible domain score}} \times \text{Actual score}$$

The attractive properties of the OES have meant it has been widely adopted. Nevertheless, Younis *et al* reported asymptomatic Oxford shoulder scores as significantly different from the presumed perfect score of 48.⁶ They suggested that PROMs can only be used effectively if the background population range is known so that postoperative improvements can be forecast more accurately.

Furthermore, Litt and Bolton-Maggs noted a trend for these scores to worsen with age and suggested that each

unit 'should identify its own baseline for the population it serves'.⁷ This was, however, a small study where the minimum sample size as indicated by the power analysis was not achieved.

The aim of our study was to ascertain the background OES for the South West Peninsula population. Our null hypothesis was that the asymptomatic South West Peninsula population OES was equal to 48 and there was no difference in scores in terms of age, sex or dominance.

Methods

A cross-sectional survey was performed of patients and healthcare professionals from the South West Peninsula population. Data was obtained from two separate district general hospitals. Participants were asked to complete a simple demographic questionnaire and an OES for each elbow. Each questionnaire was anonymous. Data were compiled on an Excel® spreadsheet (Microsoft, Redmond, WA, US) on a secure system that was password protected. Inclusion criteria consisted of participants aged 11–80 years with asymptomatic elbows while exclusion criteria consisted of participants with a history of elbow surgery, elbow injury, chronic elbow problems, an incomplete dataset and those aged less than 11 or greater than 80 years.

Data analysis was performed using SPSS® version 17 (SPSS, Chicago, IL, US) by the statistics department of the hospital. This study was approved prospectively by the hospital research, audit and ethics department.

Results

A total of 1,765 respondents (3,530 elbows) completed the survey. However, 567 questionnaires were excluded (incomplete datasets [$n=307$], pre-existing elbow pathology [$n=243$], age criteria [$n=17$]), leaving 1,198 individuals (2,396 elbows). This meant there were 701 female patients (1,402 elbows) and 497 male patients (994 elbows) available for analysis. Table 1 shows the results of the respondents by age. The median score for the whole sample population was 48 (range: 12–48, interquartile range [IQR]: 48–48). The median scores for each sampled age decade were 48.

The results were similar when comparing the scores between the sexes as well as between non-dominant and dominant elbows. Patients who considered themselves ambidextrous were considered to have both elbows dominant for the purposes of analysis. The median scores for both sexes were 48 (male range: 30–48, IQR: 48–48; female range: 12–48, IQR: 48–48). The median score for dominant elbows was 48 (range: 12–48, IQR: 48–48) and 48 (range: 19–48, IQR: 48–48) for non-dominant elbows (Tables 2 and 3).

Discussion

A perfect outcome measure must comply with a number of factors. It is required to be internally consistent so that the items being measured measure a single underlying concept. In addition, it needs to be valid (ie the extent to which the test measures what it intends to measure).⁸ It should be

Age	Number of respondents	Median	Range	Interquartile range	Mean
11–20	134	48	41–48	48–48	47.888
21–30	328	48	37–48	48–48	47.741
31–40	534	48	26–48	48–48	47.610
41–50	458	48	16–48	48–48	47.153
51–60	376	48	12–48	48–48	46.777
61–70	408	48	38–48	48–48	47.841
71–80	158	48	32–48	48–48	47.127
11–80	2,396	48	12–48	48–48	47.433

Age	Number of respondents	Median	Range	Interquartile range	Mean
11–20	69	48	44–48	48–48	47.928
21–30	168	48	38–48	48–48	47.744
31–40	269	48	26–48	48–48	47.543
41–50	234	48	16–48	48–48	47.026
51–60	198	48	12–48	48–48	46.737
61–70	210	48	38–48	48–48	47.748
71–80	81	48	32–48	48–48	46.889
11–80	1,229	48	12–48	48–48	47.356

Age	Number of respondents	Median	Range	Interquartile range	Mean
11–20	65	48	41–48	48–48	47.846
21–30	160	48	37–48	48–48	47.738
31–40	265	48	28–48	48–48	47.679
41–50	224	48	20–48	48–48	47.286
51–60	178	48	19–48	48–48	46.820
61–70	198	48	41–48	48–48	47.939
71–80	77	48	38–48	48–48	47.377
11–80	1,167	48	19–48	48–48	47.514

reliable and responsive⁹ in terms of the test revealing the same result if repeated on more than one occasion and the ability of the measure to detect true change in a patient's status over time. Finally, the perfect outcome measure should be short and practical to use.

The OES has been proven to be valid, reliable and sensitive to change after rigorous testing.⁵ It has also been shown to have high correlation with other outcome measures (Disabilities of the Arm, Shoulder and Hand questionnaire; Mayo elbow performance score; and SF-36® questionnaire).

Furthermore, the OES has numerous other qualities that have caused it to be widely adopted. It can be completed anywhere and delivered by post, making follow-up of large numbers of patients much more feasible (and cheaper) than conducting clinical assessments, its scores are easy to interpret and its application can be extended to the assessment of other (non-surgical) forms of therapy (eg physiotherapy).

A weakness of the study is the wide range of scores seen. The majority of scores (98.8%) were above 40 but a group of results were very low (there were 29 participants [1.2%] whose score was <40) even though the participants had stated on the questionnaire that they had no elbow problems, previous surgery etc. This was investigated and thought to be a set of results recorded in a rheumatology clinic that occurred unexpectedly in one of our normal data gathering orthopaedic clinics over a two-day period. Despite this, we kept to our inclusion and exclusion criteria, and these results were therefore included in our data analysis.

The developers of the OES have also published further research looking at the minimal clinically important difference (MCID) of the OES.¹⁰ The MCID is defined as the smallest change in the measure that patients perceive as meaningful and that would cause clinicians to consider revising a patient's management. It acts as a guide for clinical relevance so that changes in scores exceeding it can be assumed to be clinically significant. Any changes falling below this value, even if statistically significant, indicate changes of no clinical relevance.

Using the metric score out of 100, it was reported that a change in score of approximately 10 for the OES function scale and 18 for the OES pain and social-psychological scales is required to have a MCID.¹⁰ Extrapolation of these data using a scale of 0–48 would indicate a requirement of a change in score of 1.6 for the OES function scale and 2.88 for the OES pain and social-psychological scales to have a MCID. Combined with the fact that the majority of scores in our study were found to be very close to the perfect score of 48 and our belief that this study demonstrates the largest sample size of any published data on asymptomatic

scores encompassing almost all age groups, this extrapolation means we can suggest there was no clinical significant difference in the OESs for age, sex or hand dominance.

Conclusions

A 'normal' OES (ie a score of 48) should be used as the benchmark when assessing outcomes after elbow surgery. This is independent of age, sex and hand dominance.

Acknowledgement

We are grateful to Mrs Jo Palmer of the statistics department for all her work and advice on this paper.

References

1. Buchbinder R, Green S, Bell S *et al.* Surgery for lateral elbow pain. *Cochrane Database Syst Rev* 2002; **1**: CD003525.
2. Pynsent P, Fairbank J, Carr A. *Outcome Measures in Orthopaedics*. Oxford: Butterworth-Heinemann; 1994.
3. Jenkinson C. Measuring Health and Medical Outcomes: An Overview. In: Jenkinson C. *Measuring Health and Medical Outcomes*. London: UCL Press; 1994. pp1–6.
4. Haywood KL. Patient-reported outcome: measuring what matters or just another paper exercise? *Musculoskeletal Care* 2006; **4**: 63–66.
5. Dawson J, Doll H, Boller I *et al.* The development and validation of a patient-reported questionnaire to assess outcomes of elbow surgery. *J Bone Joint Surg Br* 2008; **90**: 466–473.
6. Younis F, Sultan J, Dix S, Hughes PJ. The range of the Oxford Shoulder Score in the asymptomatic population: a marker for post-operative improvement. *Ann R Coll Surg Engl* 2011; **93**: 629–633.
7. Litt E, Bolton-Maggs BG. *Oxford scores of the shoulder and elbow in the 'normal' population*. Presented at: British Elbow and Shoulder Society Annual Scientific Meeting; June 2009; London.
8. Binkley JM, Stratford PW, Lott SA, Riddle DL. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. *Phys Ther* 1999; **79**: 371–383.
9. Cole B, Finch E, Gowland C, Mayo N. *Physical Rehabilitation Outcome Measures*. Ottawa: Canadian Physiotherapy Association; 1994.
10. Dawson J, Doll H, Boller I *et al.* Comparative responsiveness and minimal change for the Oxford Elbow Score following surgery. *Qual Life Res* 2008; **17**: 1,257–1,267.